

ƯỚC TÍNH CỠ MẪU VÀ PHẦN MỀM ƯỚC TÍNH CỠ MẪU

TS. Phạm Đăng Quyết*

Tóm tắt:

Các câu hỏi về cỡ mẫu là phổ biến trong nghiên cứu. Trong các giáo trình thống kê hiện nay chỉ đưa ra công thức tính cỡ mẫu cho trường hợp chọn mẫu ngẫu nhiên đơn giản. Việc tính toán cỡ mẫu đơn giản này dựa vào công thức cơ bản, không tính đến thiết kế nghiên cứu, các thuật toán thống kê sử dụng trong thiết kế. Việc tính toán đơn giản bằng máy tính thông thường thường không chính xác, chỉ có tính chất giải trình đủ cho đề mục chọn mẫu và tính cỡ mẫu trong đề cương báo cáo nghiên cứu. Bài viết này trình bày việc tính toán cỡ mẫu nâng cao dựa vào một số phần mềm tính toán cỡ mẫu tốt thông qua một nghiên cứu lâm sàng cụ thể.

Ước tính số lượng đối tượng (cỡ mẫu) cần thiết cho một nghiên cứu đóng vai trò cực kì quan trọng, vì nó có thể là yếu tố quyết định sự thành công hay thất bại của nghiên cứu. Nếu cỡ mẫu không đủ thì kết luận rút ra từ công trình nghiên cứu có độ chính xác không cao, thậm chí không thể kết luận gì được. Ngược lại, nếu cỡ mẫu quá nhiều hơn cần thiết thì nguồn lực, tiền bạc và thời gian sẽ bị hao phí. Do đó, vấn đề then chốt trước khi nghiên cứu là phải ước tính cho được cỡ mẫu vừa đủ cho mục tiêu của nghiên cứu. Nếu một nghiên cứu không có một cỡ mẫu tối ưu, ý nghĩa của các kết quả trong thực tế (sự khác biệt thực sự) có thể không được phát hiện.

Cỡ mẫu "vừa đủ" tùy thuộc vào loại hình nghiên cứu vào 3 thông số chính: (i) Phương pháp thiết kế nghiên cứu và đo lường kết quả (outcome measure); (ii) Hệ số ảnh hưởng (effect size); (iii) Sai lầm mà nhà nghiên cứu chấp nhận, cụ thể là sai lầm loại I và sai lầm loại II. Không biết (hay chưa quyết định) được thiết kế nghiên cứu và không có số liệu về 3 thông số trên thì không thể nào ước tính cỡ mẫu.

1. Thiết kế nghiên cứu và tiêu chí đo lường kết quả

1.1 Thiết kế nghiên cứu

Thông tin thứ nhất trong qui trình ước tính cỡ mẫu là phương pháp thiết kế (còn gọi là thể loại nghiên cứu), bởi vì yếu tố này có ảnh hưởng đến phương pháp phân tích thống kê và vì thế ảnh hưởng đến phương pháp ước tính cỡ mẫu. Có thể phân biệt các thể loại nghiên cứu này dựa vào hai tiêu chí: Thời gian và đặc tính. Theo đó thiết kế nghiên cứu, gồm: Nghiên cứu tại một thời điểm; nghiên cứu đối chứng; nghiên cứu xuôi thời gian.

1.2 Tiêu chí đo lường kết quả

Sau khi đã xác định thể loại nghiên cứu, nhà nghiên cứu cần phải quyết định chọn một tiêu chí chính đo lường kết quả để căn cứ vào đó ước tính cỡ mẫu.

Mặc dù có rất nhiều loại biến khác nhau, cho mục đích phân tích, chúng ta hoàn toàn có thể đơn giản hóa chúng thành 2 loại là *biến liên tục* và *biến phân loại*. Biến liên tục bao gồm các biến liên tục thực sự và biến sử dụng thang đo khoảng. Trong khi đó biến phân loại bao gồm biến nhị thức, biến định

* Hội Thống kê Việt Nam

danh và biến thứ bậc. Biến số sử dụng là liên tục hay phân loại sẽ quyết định cách chúng ta tóm tắt số liệu (*trung bình* hay *tỷ lệ*), hoặc cách chúng ta lựa chọn kiểm định thống kê.

Quyết định chọn tiêu chí đo lường kết quả là một quyết định vừa mang tính thực tế, vừa mang tính khoa học. Mặt khác, tiêu chí phải đáp ứng các tiêu chuẩn khoa học về độ tin cậy và độ chính xác. Do vậy việc chọn tiêu chí đo lường kết quả nghiên cứu cần phải cân nhắc cẩn thận. Quyết định chọn tiêu chí đo lường kết quả là một quyết định quan trọng, bởi vì nó có ảnh hưởng đến cỡ mẫu rất lớn.

2. Hệ số ảnh hưởng (effect size)

Trong nghiên cứu so sánh các can thiệp khác nhau, hệ số ảnh hưởng là độ lớn của sự khác biệt giữa các nhóm. Hệ số ảnh hưởng tuyệt đối là sự khác biệt giữa các trung bình, hoặc trung bình các kết quả trong hai nhóm can thiệp khác nhau.

Hệ số ảnh hưởng, nói một cách đơn giản, là một chỉ số về độ ảnh hưởng của một thuật can thiệp. Vì phản ánh mức độ khác biệt, hệ số ảnh hưởng cho phép chúng ta tránh khỏi cách diễn dịch giới hạn bởi ngôn ngữ nhị thức (như "có hay không có ảnh hưởng?"), và tập trung vào một cách diễn dịch mang tính khoa học hơn (như "mức độ ảnh hưởng cao hay thấp cỡ nào?").

Công thức chung cho ước tính hệ số ảnh hưởng (viết tắt ES) là:

$$ES = \frac{\bar{x}_1 - \bar{x}_0}{s_0} \quad [1]$$

Trong đó:

\bar{x}_1 : Số trung bình của nhóm can thiệp;

\bar{x}_0 : Số trung bình của nhóm đối chứng;

s_0 : Độ lệch chuẩn của nhóm đối chứng.

Nên nhớ rằng độ lệch chuẩn có cùng đơn vị đo lường với độ ảnh hưởng trung bình, cho nên hệ số ảnh hưởng không có đơn vị.

Hệ số ảnh hưởng là phát hiện chính của nghiên cứu định lượng. Trong khi giá trị P có thể thông báo liệu một ảnh hưởng tồn tại, giá trị P sẽ không tiết lộ mức độ của ảnh hưởng. Trong báo cáo và nghiên cứu điển giải, cả ý nghĩa thực chất (hệ số ảnh hưởng) và ý nghĩa thống kê (giá trị P) là kết quả cần được báo cáo.

3. Sai lầm loại I, II và khái niệm về hiệu năng (power)

Trong ngữ cảnh của ước tính cỡ mẫu, các tham số thường là trung bình hoặc tỷ lệ của biến kết quả quan tâm quan trọng. Giả thiết không (H_0) là trạng thái giả thiết giữ nguyên hiện trạng, không có sự khác biệt, không có ảnh hưởng, hoặc không thay đổi. Giả thiết đảo (H_a) thường được gọi là đối thiết hoặc giả thiết thay thế, thể hiện một sự khác biệt giữa các nhóm, một ảnh hưởng thực sự, và sự từ bỏ nguyên trạng.

Chúng ta sử dụng các phương pháp kiểm định thống kê như kiểm định t , F , z , χ^2 , v.v... để đánh giá khả năng của giả thiết. Kết quả của một kiểm định thống kê có thể đơn giản chia thành hai giá trị: *Có ý nghĩa thống kê* (statistical significance), hoặc là *không có ý nghĩa thống kê* (non-significance). Có ý nghĩa thống kê ở đây thường dựa vào trị số P. Thông thường, nếu $P < 0,05$, chúng ta phát biểu kết quả có ý nghĩa thống kê, nếu $P > 0,05$ chúng ta nói kết quả không có ý nghĩa thống kê.

Trong kiểm định giả thuyết, dựa trên kết quả điều tra mẫu để đưa ra kết luận về một trong hai giả thiết. Kết luận có thể là đúng hay sai và các nhà thống kê đã phân loại cách thức mà kết luận có thể đúng hoặc

không đúng. Sai lầm trong kết luận tưởng tượng được gọi là sai lầm loại I hoặc loại II. Nếu chúng ta bác bỏ H_0 khi trong thực tế H_0

đúng, đó là sai lầm loại I. Nếu chúng ta chấp nhận H_0 khi trong thực tế H_0 sai, đó là sai lầm loại II.

Bảng 1: Các khả năng có thể xảy ra của bất kỳ kiểm định giả thuyết

	Quyết định của các nhà nghiên cứu		
		Chấp nhận H_0	Bác bỏ H_0
Thực tế của giả thuyết	Ho đúng	Đúng	Sai lầm loại I hoặc sai số α
	Ho sai	Sai lầm loại II hoặc sai số β	Đúng Hiệu năng = $1 - \beta$

Hiệu năng (power) là xác suất bác bỏ một cách chính xác H_0 . Ví dụ, kết luận rằng có một sự khác biệt khi trên thực tế, thực sự là có sự khác biệt. Tính toán cỡ mẫu thường được gọi là tính toán hiệu năng. Hiệu năng càng lớn thì cỡ mẫu càng lớn.

4. Phương pháp ước tính cỡ mẫu

Như đã đề cập trong phần đầu, để ước tính cỡ mẫu cần thiết cho một công trình nghiên cứu, ngoài thể loại nghiên cứu, chúng ta cần phải có 3 số liệu: Xác suất sai lầm loại I, hiệu năng, và hệ số ảnh hưởng. Cỡ mẫu là hàm số của ba thông số này. Gọi n là cỡ mẫu cần thiết, α là sai số loại I, β là sai số loại II (hiệu năng là $1-\beta$), hệ số ảnh hưởng là ES , thì công thức chung để ước tính cỡ mẫu là:

$$n = \frac{(z_{\alpha/2} + z_{\beta})^2}{(ES)^2}$$

Trong đó, $z_{\alpha/2}$ và z_{β} là những hằng số (thật ra là số độ lệch chuẩn) từ phân phối chuẩn (standardized normal distribution) cho xác suất sai số α và β . Bởi vì, trong công thức trên ES là mẫu số, cho nên nếu ES thấp thì số lượng cỡ mẫu sẽ tăng; ngược lại, nếu ES cao thì số lượng cỡ mẫu sẽ giảm.

Trong thực tế, ước tính hệ số ảnh hưởng thường cần thiết trước khi bắt đầu n

lực nghiên cứu, để tính toán cỡ mẫu có thể được yêu cầu để tránh sai lầm loại II. Nói cách khác, bạn phải xác định cỡ mẫu trong nghiên cứu sẽ là đủ để đảm bảo (với một mức độ cụ thể của sự chắc chắn) rằng nghiên cứu có hiệu năng chấp nhận được để ủng hộ H_0 . Đó là, nếu không có sự khác biệt được tìm thấy giữa các nhóm, thì đây là một phát hiện đúng sự thật.

Vì ảnh hưởng như thế, hệ số ảnh hưởng phải được giả định trước khi tính toán. Đây là thông số không phải lúc nào cũng có sẵn, cho nên nhà nghiên cứu cần phải xem xét các nghiên cứu trước hay độ ảnh hưởng có ý nghĩa thực tế để tính toán cỡ mẫu. Về xác suất sai số, thông thường một nghiên cứu chấp nhận sai số loại I khoảng 1% hay 5% (tức $\alpha = 0,01$ hay $0,05$), và xác suất sai số loại II khoảng $\beta = 0,1$ đến $\beta = 0,2$ (tức hiệu năng phải từ $0,8$ đến $0,9$). Mỗi trường hợp gắn liền với một hằng số $z_{\alpha/2}$ và z_{β} như vừa đề cập.

Hai hằng số này có thể viết gọn bằng công thức $C = (z_{\alpha/2} + z_{\beta})^2$. C được xác định bởi phân phối chuẩn như trình bày trong Bảng 2 dưới đây. Chẳng hạn như nếu muốn $\alpha=0,05$ và hiệu năng = $0,8$, thì hằng số C là $7,85$.

Bảng 2: Hằng số C liên quan đến sai số loại I và II

α	$\beta = 0,2$ (Hiệu năng = 0,8)	$\beta = 0,1$ (Hiệu năng = 0,9)	$\beta = 0,05$ (Hiệu năng = 0,95)
0,1	6,15	8,53	10,79
0,05	7,85	10,51	13,00
0,01	13,33	16,74	19,84

5. Hiệu ứng thiết kế (Design effect - DEFF)

Các công thức tính cỡ mẫu dựa trên phương pháp chọn mẫu ngẫu nhiên giản đơn (simple random) cho một giai đoạn. Với các nghiên cứu dùng các phương pháp chọn mẫu nhiều giai đoạn, để hiệu chỉnh cho sự khác biệt giữa thiết kế lựa chọn và chọn mẫu ngẫu nhiên giản đơn, DEFF được sử dụng để tính cỡ mẫu. DEFF là tỷ lệ giữa phương sai khi dùng cách chọn mẫu trong thiết kế lựa chọn với phương sai khi dùng phương pháp ngẫu nhiên giản đơn. DEFF=3 có nghĩa phương sai mẫu lớn gấp 3 lần phương sai mẫu nếu dùng chọn mẫu ngẫu nhiên giản đơn. DEFF được tính dựa trên kết quả nghiên cứu tương tự đã làm trên quần thể đó, nếu không có nghiên cứu tương tự thì DEFF được ước tính.

Với chọn mẫu 2 giai đoạn - thường sử dụng trong khảo sát hộ gia đình hoặc điều tra dịch tễ, DEFF=2.

Như vậy cỡ mẫu thực sự cần của các nghiên cứu sẽ bằng cỡ mẫu tính cho chọn mẫu ngẫu nhiên đơn giản nhân với DEFF. Trong nghiên cứu điều tra chọn mẫu 2 giai đoạn cỡ mẫu thường được nhân với 2 (DEFF=2).

Làm thế nào để giảm DEFF: Tăng số lượng cụm/chùm; giảm số cá thể chọn tại các cụm; số lượng cá thể chọn tại các cụm bằng nhau; sử dụng phương pháp chọn mẫu ngẫu

nhiên hệ thống để chọn đối tượng tại giai đoạn cuối của chọn mẫu.

Các yếu tố cần nhắc trước khi tính cỡ mẫu: Xác định loại số liệu (ước tính hay so sánh tỷ lệ, tỷ suất, giá trị trung bình, độ lệch chuẩn); xác định sử dụng loại kiểm định thống kê; sử dụng kiểm định thống kê so sánh một chiều hay hai chiều; xác định các trị số cho hiệu năng mẫu (power), độ chính xác tuyệt đối, khác biệt; lựa chọn công thức tính cỡ mẫu; xác định DEFF; lựa chọn phương pháp tính cỡ mẫu - tính tay hoặc sử dụng phần mềm ước tính cỡ mẫu.

6. Các phần mềm ước tính cỡ mẫu

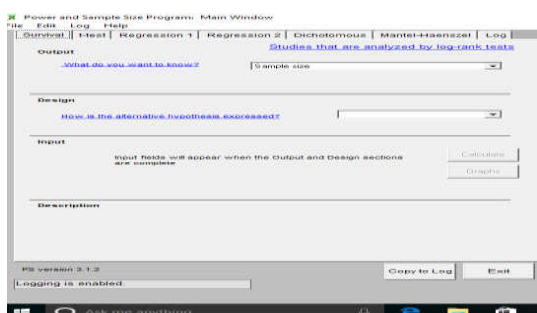
Có rất nhiều gói phần mềm để thực hiện tính toán cỡ mẫu/ hiệu năng, ví dụ: OpenEpi, PS-Power and Sample Size Calculation; EpiCalc 2000; PASS-Power Analysis and Sample Size¹¹ (commercial software); Stata (commercial software); Web-based calculator... PS có lẽ là lựa chọn phần mềm tốt nhất, vì phần mềm này sử dụng thiết kế nghiên cứu thông dụng nhất, tương đối dễ sử dụng và quan trọng là tải về dễ dàng và miễn phí. Ngoài ra, người dùng có thể khám phá sở thích của cá nhân cho riêng mình.

6.1 Phần mềm ước tính cỡ mẫu PS

PS là một chương trình máy tính tương tác để thực hiện tính toán hiệu năng và cỡ mẫu. Chương trình PS có thể được sử dụng cho nghiên cứu với các đo lường nhị thức, liên tục, hoặc câu trả lời về sống còn. Người dùng chỉ rõ các giả thiết thay thế về tỷ lệ trả lời khác nhau, trung bình, thời gian sống, nguy cơ tương đối, hoặc tỷ lệ nguy cơ. Các thiết kế nghiên cứu độc lập hay cặp đôi (phù hợp) có thể được sử dụng. Hiệu năng, cỡ mẫu, và các giả thiết thay thế phát hiện được là quan hệ với nhau. Người dùng chỉ rõ hai trong ba đại

lượng này và chương trình thu được đại lượng thứ ba. Mô tả mỗi tính toán, được viết bằng tiếng Anh, được tạo ra và có thể sao chép được vào các tài liệu của người dùng. Hỗ trợ tương tác là có sẵn. Chương trình cung cấp các phương pháp phù hợp cho kiểm định t cặp đôi và độc lập, phân tích sống còn (Survival Analysis, phương pháp rất hay gặp trong nghiên cứu lâm sàng), nghiên cứu phù hợp và không phù hợp về các sự kiện nhị thức, kiểm định Mantel-Haenszel, và hồi quy tuyến tính. Chương trình có thể tạo ra các đồ thị của các mối quan hệ giữa hiệu năng, cỡ mẫu và các giả thiết thay thế phát hiện được. Nó có thể vẽ đồ thị của bất kỳ hai trong các đại lượng này trong khi duy trì đại lượng thứ ba không đổi. Các trục tuyến tính hoặc logarit có thể được sử dụng và nhiều đường cong có thể được vẽ trên mỗi đồ thị. Đồ thị có thể được sao chép và dán vào tài liệu khác hoặc các chương trình để chỉnh sửa thêm.

Hình 1: Giao diện chính Chương trình PS



(1) Ước tính cỡ mẫu cho một chỉ tiêu trung bình

Ví dụ: Tính cỡ mẫu để ước tính chiều cao đàn ông Việt với sai số δ trong vòng 1 cm. Biết rằng độ lệch chuẩn trong các nghiên cứu trước đây là 4,6 cm với khoảng tin cậy 95%, hay sai số $\alpha=0,05$ và sai số $\beta=0,2$ (hiệu năng mẫu=0,8).

Để tính cỡ mẫu cần thiết ta làm như sau: Từ thanh công cụ trên màn hình chọn t-

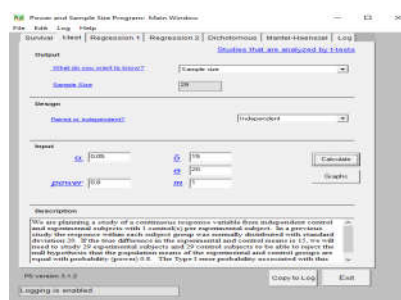
test, hộp Output chọn Sample size, hộp Design chọn Paired, Input $\alpha=0,05$, power=0,8, $\sigma=4,6$, $\delta=1$, nhấp chuột lên Calculate. Màn hình sẽ hiển thị hộp Sample Size với kết quả là 168 cm.

Nói cách khác, chúng ta cần phải đo chiều cao ở 168 đối tượng để ước tính chiều cao đàn ông Việt với sai số trong vòng 1 cm. Nếu sai số chấp nhận là 0,5 cm (thay vì 1 cm), số lượng đối tượng cần thiết là 666. Qua các ước tính này, chúng ta dễ dàng thấy cỡ mẫu tùy thuộc rất lớn vào độ sai số mà chúng ta chấp nhận. Muốn có ước tính càng chính xác, chúng ta cần càng nhiều đối tượng nghiên cứu.

(2) Ước tính cỡ mẫu cho so sánh hai nhóm độc lập

Một nghiên cứu đã đề xuất mong muốn tìm hiểu ảnh hưởng của một loại thuốc tăng huyết áp mới (nhóm can thiệp) so với điều trị thông thường (nhóm đối chứng). Các nghiên cứu trước cho thấy rằng tối thiểu khác biệt lâm sàng quan trọng là 15 mmHg và độ lệch chuẩn gộp (σ) là 20 mmHg.

Sử dụng phần mềm PS (hình bên) người ta ước tính rằng 29 đối tượng



sẽ là cần thiết trong mỗi nhóm đối chứng và nhóm can thiệp (mức $\alpha = 0,05$, hiệu năng 80%) để phát hiện một sự khác biệt có ý nghĩa thống kê trong huyết áp trung bình giữa hai nhóm (nếu tồn tại). Tăng khác biệt trung bình đến 20 mmHg sẽ yêu cầu 17 người tham gia mỗi nhóm, tức là giảm đáng kể cỡ mẫu. Giảm σ đến 10 mmHg sẽ yêu cầu chỉ có 5 người tham gia trong mỗi nhóm.

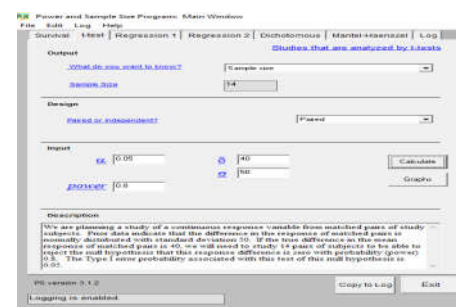
(3) Ước tính cỡ mẫu cho so sánh hai tỷ lệ (Chi-square test)

Một can thiệp nâng cao sức khỏe để giảm hút thuốc sẽ được giới thiệu trong một đất nước với tỷ lệ hiện hành hút thuốc cao. Tỷ lệ hút thuốc hiện tại trong một báo cáo công bố gần đây là 65% và ước tính rằng sự can thiệp sẽ làm giảm mức độ hút thuốc gần 30%. Một cỡ mẫu là 42 trong mỗi nhóm được yêu cầu ($P = 0,05$, hiệu năng 80%). Tăng mức hiệu năng tới 90% đòi hỏi cỡ mẫu 56 cho mỗi nhóm.



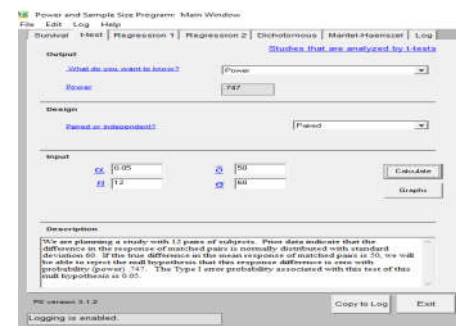
(4) Ước tính cỡ mẫu cho hai nhóm cặp đôi (paired t-test)

Có bằng chứng cho thấy clofibrate thay đổi giá trị trung bình mức cholesterol? Cholesterol được đo trước và sau khi nhận được clofibrate. Từ nghiên cứu trước, một sự khác biệt trung bình 40 mg/dl được coi là lâm sàng có ý nghĩa, với độ lệch chuẩn là 50. Một cỡ mẫu 14 là cần thiết ($P = 0,05$, hiệu năng 80%), nếu $P = 0,01$ đòi hỏi 21 đối tượng, tức là cỡ mẫu tăng lên.



(5) Tính toán hiệu năng sau nghiên cứu

Nghiên cứu cholesterol đề cập trên đã được tiến hành, nhưng do kết quả tỷ lệ trả lời/ bỏ cuộc không cao, chỉ có 12 bệnh nhân được tuyển chọn. Khác biệt trung bình là 50 và σ là 60 mg/dl được tìm thấy. Hiệu năng của nghiên cứu này là bao nhiêu? Thực hiện một phân tích hồi cứu cho hay hiệu năng là 75%, có nghĩa là nghiên cứu là dưới mức hiệu năng chấp nhận được vì nó nhỏ hơn 80%.

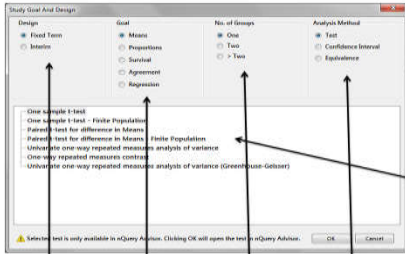


6.2 Phần mềm ước tính cỡ mẫu chuyên nghiệp

Phần mềm tính toán cỡ mẫu chuyên nghiệp hiện nay như SAS, hoặc n-query. Tuy nhiên những phần mềm này rất đắt (cỡ 1 triệu USD cho SAS, 1050 USD cho n-query).

Nghiên cứu viên nhập thiết kế nghiên cứu, thuật toán thống kê sử dụng và các dữ liệu liên quan đến biến số vào phần mềm để tính ra cỡ mẫu. Nghiên cứu viên có thể phân tích độ nhạy mỗi liên quan giữa hiệu năng và cỡ mẫu. Nghiên cứu viên phải có kiến thức chuyên sâu về thiết kế nghiên cứu và các thuật toán thống kê để có thể sử dụng được phần mềm cũng như phân giải kết quả.

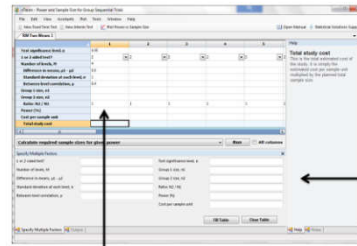
Dưới đây là hình ảnh 4 bước sử dụng phần mềm n-query tính toán cỡ mẫu:



Bước 1: Nhập dữ liệu về thiết kế nghiên cứu, tác động và thuật toán thống kê phân tích

Cụ thể hóa nhập liệu thuật toán thống kê

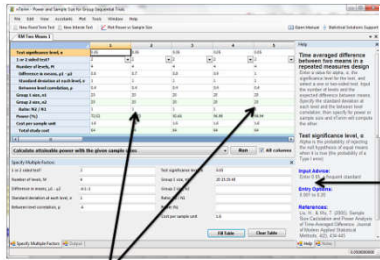
Loại thiết kế Loại dữ liệu đầu ra Số nhóm mẫu Loại phân tích thống kê



Bước 2: Xác định giá trị cụ thể của biến số và mong muốn đặc tính mẫu

Các ô tính ra cỡ mẫu

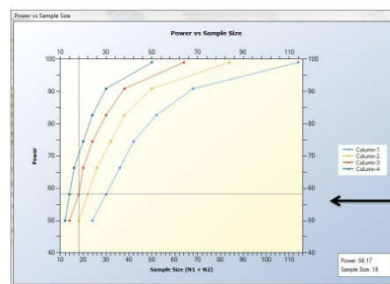
Nhập giá trị cụ thể của biến số và mong muốn đặc tính mẫu



Bước 3: Xác định hiệu năng mẫu từ các cỡ mẫu và ngược lại

Các ô kết quả

Nhập giá trị cụ thể của biến số và mong muốn đặc tính mẫu



Bước 4: Phân tích độ nhạy của mối quan hệ Hiệu năng mẫu - cỡ mẫu

Biểu đồ của phân tích độ nhạy

Kết luận

Như đã đề cập ở trên, cách tính cỡ mẫu chỉ cho chúng ta một ước lượng thô của cỡ mẫu cần thiết bởi vì nó dựa trên sự ước đoán về giá trị của thông số, quyết định chủ quan của chúng ta về hậu quả mà chúng ta muốn phát hiện và công thức được sử dụng là công thức gần đúng. Do đó con số tính ra giúp chúng ta phân biệt giữa cỡ mẫu 50 và 100 chứ không phân biệt cỡ mẫu 50 và 53. Chúng ta phải cân đối giữa điều chúng ta mong muốn và tính khả thi. Đôi khi có thể dùng công thức tính cỡ mẫu để đi ngược lại hiệu năng của nghiên cứu. Thí dụ nếu chúng ta có một kinh phí hạn chế để thực hiện một nghiên cứu nên chỉ có một cỡ mẫu nhất định. Chúng ta có thể tính ngược lại từ cỡ mẫu để biết hiệu năng của nghiên cứu. Nếu hóa ra hiệu năng của nghiên cứu rất thấp (thí dụ như 20%) tốt nhất chúng ta không nên tiến hành nghiên cứu vì chúng ta đã nắm chắc kết quả thất bại. Nếu một nghiên cứu có nhiều mục tiêu thì cỡ mẫu đủ cho một mục tiêu này có thể không đủ cho

mục tiêu khác. Để tính cỡ mẫu, tốt nhất phải chú trọng vào biến số (hoặc những biến số quan trọng nhất). Tính cỡ mẫu không khó, cái khó là phải cung cấp những giá trị giả định của nghiên cứu: Sai lầm loại I, hiệu năng, năng lực nghiên cứu, sự khác biệt mà chúng ta muốn phát hiện.

Tài liệu tham khảo:

1. Đại học Y tế công cộng (2004), *Thống kê y tế II - Phân tích số liệu*, Bộ môn Thống kê - Tin học, Hà Nội;
2. Evie McCrum-Gardner, "Sample size and power calculations made simple", *International Journal of Therapy and Rehabilitation*, January 2010, Vol 17, No 1;
3. Nguyễn Văn Tuấn, *Phương pháp ước tính cỡ mẫu cho một nghiên cứu y học*, www.ykhoa.net/baigiang/lamsangthongke/list_k_uoctinhcomau.pdf;
4. Nguyễn Trương Nam, *Xác định cỡ mẫu nghiên cứu*, Viện nghiên cứu Y xã hội học, thongke.info.vn/Download.aspx/.../1/ISMS_Ti_nhtoan_CoMau.pdf.